Inequality for p-norms of positive matrices

Christopher King

Department of Mathematics
Northeastern University
Boston MA 02115
king@neu.edu

February 11, 2003

Abstract

This paper derives an inequality relating the p-norm of a positive 2×2 block matrix to the p-norm of the 2×2 matrix obtained by replacing each block by its p-norm. The inequality had been known for integer values of p, so the main contribution here is the extension to all values $p \geq 1$. In a special case the result reproduces Hanner's inequality. As an application in quantum information theory, the inequality is used to obtain some results concerning maximal p-norms of product channels.

1 Introduction and statement of results

Quantum information theory has raised some interesting mathematical questions about completely positive trace preserving maps. Such maps describe the evolution of open quantum systems, or quantum systems in the presence of noise [3]. Many of these questions are related to the quantum entropy of states, and the associated notion of the p-norm of a state. In one case [6] the investigation of the additivity question for product channels (which will be explained in Section 4) led to an inequality for p-norms of 2×2 block matrices for integer values of p. The present paper is devoted to showing that this inequality extends to non-integer values of p. The inequality turns out to be closely related to Hanner's inequality, which was proved for all $p \geq 1$ by Ball, Carlen and Lieb [2], and the method of proof in this paper uses many of the ideas and results from that paper.

Let M be a $2n \times 2n$ positive semi-definite matrix. It can be written in the block form

$$M = \begin{pmatrix} X & Y \\ Y^* & Z \end{pmatrix} \tag{1}$$

where X, Y, Z are $n \times n$ matrices. The condition $M \geq 0$ requires that $X \geq 0$ and $Z \geq 0$, and also that $Y = X^{1/2} R Z^{1/2}$ where R is a contraction.

Recall that the p-norm of a matrix A is defined as

$$||A||_p = \Big(\mathrm{Tr}(A^* A)^{p/2} \Big)^{1/p} \tag{2}$$

Define the 2×2 matrix

$$m = \begin{pmatrix} ||X||_p & ||Y||_p \\ ||Y||_p & ||Z||_p \end{pmatrix} \tag{3}$$

From Hölder's inequality it follows that

$$||Y||_p = ||X^{1/2}\, R Z^{1/2}||_p \leq ||X||_p^{1/2}\ ||Z||_p^{1/2} \tag{4}$$

which implies that $m \geq 0$ also.

Theorem 1 *The following inequalities hold:*

a) *for* $1 \leq p \leq 2$,

$$||M||_p \geq ||m||_p \tag{5}$$

b) *for* $2 \leq p \leq \infty$,

$$||M||_p \leq ||m||_p \tag{6}$$

Theorem 1 is easily proved for integer values of p using Hölder's inequality (see [6] for details). In the case where $X = Z$ and $Y = Y^*$, the norms of M and m simplify in the following way:

$$\begin{aligned} ||M||_p^p &= ||X+Y||_p^p + ||X-Y||_p^p && (7) \\ ||m||_p^p &= \Big(||X||_p + ||Y||_p\Big)^p + \Big|||X||_p - ||Y||_p\Big|^p && (8) \end{aligned}$$

With these substitutions, the inequalities (5) and (6) are seen to be special cases of Hanner's inequality [5] for the matrix spaces C_p, which are the non-commutative versions of the function spaces L_p. Hanner's inequality for C_p was proved by Ball, Carlen and Lieb [2]. One of the other main results of the paper [2] is the 2-uniform convexity (with best constant) of the space C_p for $1 < p \leq 2$, which is expressed by the inequality

$$\left(\frac{||X+Y||_p^p + ||X-Y||_p^p}{2}\right)^{2/p} \geq ||X||_p^2 + (p-1)||Y||_p^2 \tag{9}$$

This inequality can be re-expressed in terms of the matrices M and m as follows:

$$||M||_p \geq 2^{1/p-1}\Big((2-p)\ (\mathrm{Tr}\ m)^2 + (2p-2)\ \mathrm{Tr}(m^2)\Big)^{1/2} \tag{10}$$

Using Gross's two-point inequality [4] and Theorem 1, it can be easily shown that the inequality (10) is also valid in the general case where M and m are given by (1) and (3).

The proof of Theorem 1 has three main ingredients: for convenience we state them as separate lemmas here. The first ingredient is a slight modification of a convexity result from [2].

Lemma 2 *Let* $M = \begin{pmatrix} X & Y \\ Y^* & Z \end{pmatrix} \geq 0$ *where* X, Y, Z *are* $n \times n$ *matrices. For fixed* Y*, and for* $1 \leq p \leq 2$*, the function*

$$(X, Z) \longmapsto \mathrm{Tr} M^p - \mathrm{Tr} X^p - \mathrm{Tr} Z^p \tag{11}$$

is jointly convex in X *and* Z*.*

The second ingredient extends a convexity result of Hanner [5] to the case of positive 2×2 matrices with positive coefficients.

Lemma 3 *Let* $A = \begin{pmatrix} a & c \\ c & b \end{pmatrix} > 0$ *where* $a, b, c \geq 0$*. For* $1 \leq p \leq 2$*, the function*

$$g(A) = \mathrm{Tr} \begin{pmatrix} a^{1/p} & c^{1/p} \\ c^{1/p} & b^{1/p} \end{pmatrix}^p \tag{12}$$

is convex in A*.*

The third ingredient is a monotonicity result for positive 2×2 matrices.

Lemma 4 *Let* $A = \begin{pmatrix} a & c \\ c & b \end{pmatrix} > 0$ *where* $a, b, c \geq 0$*. For fixed* c*, and for* $1 \leq p \leq 2$*, the function*

$$(a, b) \longmapsto \mathrm{Tr} A^p - a^p - b^p \tag{13}$$

is decreasing in a *and* b*.*

The paper is organised as follows. In Section 2 we present the proof of Theorem 1 using Lemmas 2, 3 and 4. These lemmas are proved in Section 3, and Section 4 describes an application of Theorem 1 in Quantum Information Theory.

2 Proof of Theorem 1

Many of the ideas in this proof are taken from the proof of Hanner's inequality in [2]. First, we borrow the duality argument from Section IV of that paper to show that part (b) follows from part (a). For $p \geq 2$ define $q \leq 2$ to be its conjugate index. Then there is a $2n \times 2n$ matrix K satisfying $||K||_q = 1$ such that

$$||M||_p = \sup_{L:||L||_q=1} \mathrm{Tr}(LM) = \mathrm{Tr}(KM) \tag{14}$$

The positivity of M means that K can be assumed to be positive. Let

$$K = \begin{pmatrix} A & C \\ C^* & B \end{pmatrix} \geq 0 \tag{15}$$

then

$$\begin{aligned} \mathrm{Tr}(KM) &= \mathrm{Tr}(AX) + \mathrm{Tr}(CY^*) + \mathrm{Tr}(C^*Y) + \mathrm{Tr}(BZ) \\ &\leq ||A||_q \, ||X||_p + 2||C||_q \, ||Y||_p + ||B||_q \, ||Z||_p \\ &= \mathrm{Tr} \begin{pmatrix} ||A||_q & ||C||_q \\ ||C||_q & ||B||_q \end{pmatrix} m \\ &\leq \Big|\Big| \begin{pmatrix} ||A||_q & ||C||_q \\ ||C||_q & ||B||_q \end{pmatrix} \Big|\Big|_q \, ||m||_p \\ &\leq ||K||_q \, ||m||_p \\ &= ||m||_p \end{aligned} \tag{16}$$

The first and second inequalities are applications of Hölder's inequality, the last inequality uses part (a) of Theorem 1.

Next we turn to the proof of part (a) of Theorem 1. The inequality becomes an equality at the values $p = 1, 2$, so we will assume henceforth that $1 < p < 2$. Using the singular value decomposition we can write

$$Y = UDV^* \tag{17}$$

where U, V are unitary matrices and $D \geq 0$ is diagonal. Unitary invariance of the p norm implies that

$$||M||_p = \Big|\Big| \begin{pmatrix} U^*XU & D \\ D & V^*ZV \end{pmatrix} \Big|\Big|_p \tag{18}$$

and also that $||X||_p = ||U^* X U||_p$, $||Z||_p = ||V^* Z V||_p$ and $||Y||_p = ||D||_p$. So without loss of generality we will assume henceforth that Y is diagonal and non-negative.

Next we use a diagonalization argument from Section III of [2]. Let $U_1, \dots, U_{2^n}$ denote the 2^n diagonal $n \times n$ matrices with diagonal entries ± 1. Then for any $n \times n$ matrix A we have

$$A_d = \sum_{i=1}^{2^n} 2^{-n} \; U_i A U_i^* \tag{19}$$

where A_d is the diagonal part of A. Since Y is diagonal this implies that

$$\sum_{i=1}^{2^n} 2^{-n} \begin{pmatrix} U_i & 0 \\ 0 & U_i \end{pmatrix} \begin{pmatrix} X & Y \\ Y & Z \end{pmatrix} \begin{pmatrix} U_i^* & 0 \\ 0 & U_i^* \end{pmatrix} = \begin{pmatrix} X_d & Y \\ Y & Z_d \end{pmatrix} \tag{20}$$

and by the same reasoning

$$\sum_{i=1}^{2^n} 2^{-n} \begin{pmatrix} U_i & 0 \\ 0 & U_i \end{pmatrix} \begin{pmatrix} X & 0 \\ 0 & Z \end{pmatrix} \begin{pmatrix} U_i^* & 0 \\ 0 & U_i^* \end{pmatrix} = \begin{pmatrix} X_d & 0 \\ 0 & Z_d \end{pmatrix} \tag{21}$$

Now we combine (20) and (21) with the convexity result Lemma 2, which gives

$$\mathrm{Tr} \begin{pmatrix} X & Y \\ Y & Z \end{pmatrix}^p - \mathrm{Tr} \begin{pmatrix} X & 0 \\ 0 & Z \end{pmatrix}^p \geq \mathrm{Tr} \begin{pmatrix} X_d & Y \\ Y & Z_d \end{pmatrix}^p - \mathrm{Tr} \begin{pmatrix} X_d & 0 \\ 0 & Z_d \end{pmatrix}^p \tag{22}$$

The matrices X_d, Y, Z_d are all diagonal with non-negative entries. Denote these entries by $(x_1, \dots, x_n)$, $(y_1, \dots, y_n)$ and $(z_1, \dots, z_n)$ respectively. Then

$$\mathrm{Tr} \begin{pmatrix} X_d & Y \\ Y & Z_d \end{pmatrix}^p = \sum_{i=1}^{n} \mathrm{Tr} \begin{pmatrix} x_i & y_i \\ y_i & z_i \end{pmatrix}^p \tag{23}$$

Now for $i = 1, \dots, n$ define

$$a_i = x_i^p, \quad b_i = z_i^p, \quad c_i = y_i^p \tag{24}$$

and introduce the 2×2 matrices

$$A_i = \begin{pmatrix} a_i & c_i \\ c_i & b_i \end{pmatrix} \tag{25}$$

It follows that

$$\begin{aligned} ||X_d||_p &= (a_1 + \cdots + a_n)^{1/p} \\ ||Y||_p &= (c_1 + \cdots + c_n)^{1/p} \\ ||Z_d||_p &= (b_1 + \cdots + b_n)^{1/p} \end{aligned} \tag{26}$$

and the definition (12) implies that

$$\mathrm{Tr} \begin{pmatrix} ||X_d||_p & ||Y||_p \\ ||Y||_p & ||Z_d||_p \end{pmatrix}^p = g(A_1 + \cdots + A_n) \tag{27}$$

Furthermore (23) implies that

$$\mathrm{Tr} \begin{pmatrix} X_d & Y \\ Y & Z_d \end{pmatrix}^p = g(A_1) + \cdots + g(A_n) \tag{28}$$

Also, for any positive number k we have $g(kA) = kg(A)$. Combining this with the convexity result Lemma 3 gives

$$g(A_1 + \cdots + A_n) \leq g(A_1) + \cdots + g(A_n), \tag{29}$$

which from (28) and (27) implies that

$$\mathrm{Tr} \begin{pmatrix} X_d & Y \\ Y & Z_d \end{pmatrix}^p \geq \mathrm{Tr} \begin{pmatrix} ||X_d||_p & ||Y||_p \\ ||Y||_p & ||Z_d||_p \end{pmatrix}^p \tag{30}$$

Combining (22) with (30) gives

$$\begin{aligned} & \mathrm{Tr} \begin{pmatrix} X & Y \\ Y & Z \end{pmatrix}^p - \mathrm{Tr} \begin{pmatrix} X & 0 \\ 0 & Z \end{pmatrix}^p \\ \geq\ & \mathrm{Tr} \begin{pmatrix} ||X_d||_p & ||Y||_p \\ ||Y||_p & ||Z_d||_p \end{pmatrix}^p - \mathrm{Tr} \begin{pmatrix} ||X_d||_p & 0 \\ 0 & ||Z_d||_p \end{pmatrix}^p \end{aligned} \tag{31}$$

Furthermore

$$||X_d||_p \leq ||X||_p, \qquad ||Z_d||_p \leq ||Z||_p \tag{32}$$

Applying Lemma 4 to the right side of (31) shows that

$$\begin{aligned} & \mathrm{Tr} \begin{pmatrix} ||X_d||_p & ||Y||_p \\ ||Y||_p & ||Z_d||_p \end{pmatrix}^p - \mathrm{Tr} \begin{pmatrix} ||X_d||_p & 0 \\ 0 & ||Z_d||_p \end{pmatrix}^p \\ \geq\ & \mathrm{Tr} \begin{pmatrix} ||X||_p & ||Y||_p \\ ||Y||_p & ||Z||_p \end{pmatrix}^p - \mathrm{Tr} \begin{pmatrix} ||X||_p & 0 \\ 0 & ||Z||_p \end{pmatrix}^p \end{aligned} \tag{33}$$

Furthermore

$$\mathrm{Tr}\begin{pmatrix} X & 0 \\ 0 & Z \end{pmatrix}^p = \mathrm{Tr}\begin{pmatrix} ||X||_p & 0 \\ 0 & ||Z||_p \end{pmatrix}^p \tag{34}$$

and therefore (31) and (33) imply the result Theorem 1.

3 Proofs of Lemmas

3.1 Proof of Lemma 2

This result is a slight modification of a convexity result proved in Section IV of [2]. For a positive matrix $M = \begin{pmatrix} X & Y \\ Y^* & Z \end{pmatrix} \geq 0$, define $M_d = \begin{pmatrix} X & 0 \\ 0 & Z \end{pmatrix} \geq 0$ and $F = M - M_d$. Let

$$D = \begin{pmatrix} D_1 & 0 \\ 0 & D_2 \end{pmatrix} = D^* \tag{35}$$

be a block diagonal self-adjoint matrix, and define

$$\begin{aligned} \phi(s) &= \mathrm{Tr}(M + sD)^p - \mathrm{Tr}(M_d + sD)^p \\ &= \mathrm{Tr}(M_d + F + sD)^p - \mathrm{Tr}(M_d + sD)^p \end{aligned}$$

Then for $1 \leq p \leq 2$ the second derivative of ϕ has the following integral representation (see [2] for details):

$$\phi''(0) = p\gamma_p \int_0^\infty t^{p-1}\mathrm{Tr}\Big(\frac{1}{t + M_d + F} D \frac{1}{t + M_d + F} D - \frac{1}{t + M_d} D \frac{1}{t + M_d} D\Big) dt \tag{36}$$

for some constant γ_p. Furthermore, the matrices $M_d + F + sD$ and $M_d - F + sD$ have the same spectrum, hence (36) can be written

$$\begin{aligned} \phi''(0) = \frac{p}{2}\gamma_p \int_0^\infty t^{p-1}\mathrm{Tr}\Big(&\frac{1}{t + M_d + F} \quad D \quad \frac{1}{t + M_d + F} D \\ &+\frac{1}{t + M_d - F} \quad D \quad \frac{1}{t + M_d - F} D \\ &-2\frac{1}{t + M_d} \quad D \quad \frac{1}{t + M_d} D\Big) dt \end{aligned} \tag{37}$$

Ball, Carlen and Lieb [2] proved that for $t \geq 0$, and for any self-adjoint matrix A, the map

$$X \longmapsto \mathrm{Tr}\frac{1}{t+X}A\frac{1}{t+X}A \tag{38}$$

is convex on the set of positive matrices. Applying this to (37) with $X = M_d$ and $A = D$ shows that $\phi''(0) \geq 0$, which is the convexity result in Lemma 2.

3.2 Proof of Lemma 3

Since g is homogeneous it is sufficient to prove that

$$\frac{d}{dt}g(A+tB)|_{t=0} \leq g(B) \tag{39}$$

for any A, B of the specified form. Let

$$A = \begin{pmatrix} a & c \\ c & b \end{pmatrix}, \quad B = \begin{pmatrix} x & y \\ y & z \end{pmatrix} \tag{40}$$

Define

$$M = \begin{pmatrix} a^{1/p} & c^{1/p} \\ c^{1/p} & b^{1/p} \end{pmatrix}, \quad L = \begin{pmatrix} a^{(1-p)/p}x & c^{(1-p)/p}y \\ c^{(1-p)/p}y & b^{(p-1)/p}z \end{pmatrix} \tag{41}$$

Then

$$\frac{d}{dt}g(A+tB)|_{t=0} = \mathrm{Tr}M^{p-1}\,L \tag{42}$$

The idea of the proof is to maximise the right side of (42) as a function of a, b, c, and show that the maximum is achieved when A and B are proportional, in which case the bound is an equality. To this end write the spectral decomposition of M in the form

$$M = \begin{pmatrix} m_{11} & m_{12} \\ m_{12} & m_{22} \end{pmatrix} = \lambda P_1 + \mu P_2 \tag{43}$$

where P_i are projectors onto the normalised eigenvectors of M, and λ, μ are the eigenvalues. If we assume that $\lambda \geq \mu$ then for some $0 \leq t \leq 1$ we have

$$m_{11} = \lambda t + \mu(1-t) \tag{44}$$

$$m_{12} = \sqrt{t(1-t)}(\lambda - \mu) \tag{45}$$

$$m_{22} = \lambda(1-t) + \mu t \tag{46}$$

Furthermore it also follows that

$$M^{p-1} = \begin{pmatrix} k_{11} & k_{12} \\ k_{12} & k_{22} \end{pmatrix} = \lambda^{p-1} P_1 + \mu^{p-1} P_2 \tag{47}$$

where

$$k_{11} = \lambda^{p-1} t + \mu^{p-1}(1-t) \tag{48}$$
$$k_{12} = \sqrt{t(1-t)}(\lambda^{p-1} - \mu^{p-1}) \tag{49}$$
$$k_{22} = \lambda^{p-1}(1-t) + \mu^{p-1} t \tag{50}$$

Substituting into (42) gives

$$\mathrm{Tr} M^{p-1} L = k_{11} m_{11}^{1-p} x + 2k_{12} m_{12}^{1-p} y + k_{22} m_{22}^{1-p} z \tag{51}$$

Equation (51) is invariant under a rescaling of M. Define

$$h = \frac{\mu}{\lambda}, \qquad 0 \leq h \leq 1 \tag{52}$$

then (51) is a function of t and h, and can be written as

$$\mathrm{Tr} M^{p-1} L = F(t,h) = F_1(t,h)x + F_2(t,h)y + F_3(t,h)z \tag{53}$$

where

$$F_1(t,h) = \frac{t + (1-t)h^{p-1}}{(t + (1-t)h)^{p-1}} \tag{54}$$
$$F_2(t,h) = 2\Big(t(1-t)\Big)^{1-p/2} \frac{1 - h^{p-1}}{(1-h)^{p-1}} \tag{55}$$
$$F_3(t,h) = F_1(1-t,h) \tag{56}$$

The goal is to maximise $F(t,h)$ over t and h. Define

$$G = (t + (1-t)h)(1 - h^{p-1}) - (p-1)(1-h)(t + (1-t)h^{p-1}) \tag{57}$$
$$H = ((1-t) + th)(1 - h^{p-1}) - (p-1)(1-h)((1-t) + th^{p-1}) \tag{58}$$

and also let

$$\xi = x(t + (1-t)h)^{-p} \tag{59}$$
$$\eta = y(1 - t + th)^{-p} \tag{60}$$
$$\zeta = z(1-h)^{-p}\,(t(1-t))^{-p/2} \tag{61}$$

Then explicit calculation shows that

$$\frac{\partial F}{\partial t} = G\xi - H\eta - (G - H)\zeta \tag{62}$$

and

$$\frac{\partial F}{\partial h} = -t(1-t)(p-1)(1-h^{p-2})(\xi + \eta - 2\zeta) \tag{63}$$

Solving the equations

$$\frac{\partial F}{\partial t} = \frac{\partial F}{\partial h} = 0 \tag{64}$$

leads to one of the possibilities (a) $t = 0$ or $t = 1$, (b) $h = 1$, (c) $\xi = \eta = \zeta$. In both cases (a) and (b), the matrix M must be diagonal, in which case (42) becomes

$$\mathrm{Tr} M^{p-1} L = \mathrm{Tr} B = \mathrm{Tr} \begin{pmatrix} x^{1/p} & 0 \\ 0 & z^{1/p} \end{pmatrix}^{p} \leq g(B) \tag{65}$$

which establishes the result. Case (c) implies that M must be proportional to the matrix

$$\begin{pmatrix} x^{1/p} & y^{1/p} \\ y^{1/p} & z^{1/p} \end{pmatrix} \tag{66}$$

and substituting into (42) then gives

$$\mathrm{Tr} M^{p-1} L = g(B) \tag{67}$$

hence the result is proved.

3.3 Proof of Lemma 4

By the convexity result Lemma 3, it is sufficient to prove that the function $(a, b) \mapsto \mathrm{Tr} A^p - a^p - b^p$ is decreasing as $a, b \to \infty$. For $a >> 1$, and for $1 < p < 2$, easy estimates show that

$$\mathrm{Tr} A^p - a^p - b^p \simeq p c^2 a^{p-2} \tag{68}$$

which is indeed decreasing. Similarly for b.

4 Application to qubit maps

Quantum information theory has generated an interesting conjecture concerning completely positive maps on matrix algebras. Let Φ be a completely positive trace-preserving (CPTP) map on the algebra of $n \times n$ matrices. The minimal entropy of Φ is defined by

$$S_{\min}(\Phi) = \inf_{\rho} S(\Phi(\rho)) \tag{69}$$

where S is the von Neumann entropy and the inf runs over $n \times n$ density matrices (satisfying $\rho \geq 0$ and $\mathrm{Tr}\rho = 1$). Minimal entropy is conjectured to be additive for product maps, that is, it is conjectured that

$$S_{\min}(\Phi_1 \otimes \Phi_2) = S_{\min}(\Phi_1) + S_{\min}(\Phi_2) \tag{70}$$

for any pair of CPTP maps Φ_1 and Φ_2. The conjecture (70) has been established in some special cases [8], [7] but a general proof remains elusive.

For related reasons, Amosov, Holevo and Werner [1] defined the maximal p-norm for a CPTP map to be

$$\nu_p(\Phi) = \sup_{\rho} ||\Phi(\rho)||_p \tag{71}$$

where the sup runs again over density matrices. They conjectured that this quantity is multiplicative for product maps, that is

$$\nu_p(\Phi_1 \otimes \Phi_2) = \nu_p(\Phi_1)\ \nu_p(\Phi_2) \tag{72}$$

Holevo and Werner later discovered a family of counterexamples to this conjecture for $p \geq 4.79$, using maps which act on 3×3 or higher dimensional matrices [9]. The conjecture remains open if at least one of the pair is a qubit map (which acts on 2×2 matrices) or if $p \leq 4$.

As an application of Theorem 1, we now show that it implies the result (72) in one special case, namely when Φ_1 is the qubit depolarizing channel and $p \geq 2$. This result was derived previously using a lengthier argument [7], and the purpose of this presentation is to explore an alternative method which may allow new approaches to the additivity problem. Indeed, the method shown below can be easily extended to cover all unital qubit channels and even some non-unital qubit maps, thus extending the results in [6] which were derived for integer values of p. Unfortunately, the restriction to $p \geq 2$ does not allow any conclusions to be drawn about additivity of minimal entropy.

The depolarizing channel Δ acts on a state $\rho = \begin{pmatrix} a & c \\ \overline{c} & b \end{pmatrix}$ by

$$\Delta(\rho) = \lambda\rho + \frac{1-\lambda}{2} I = \begin{pmatrix} \lambda_+ a + \lambda_- b & \lambda c \\ \lambda \overline{c} & \lambda_- a + \lambda_+ b \end{pmatrix} \tag{73}$$

where λ is a real parameter and $\lambda_\pm = (1 \pm \lambda)/2$. We will suppose here that $\lambda \geq 0$ for convenience. The maximal p-norm of Δ is easily computed to be

$$\nu_p(\Delta) = \left(\left(\frac{1+\lambda}{2}\right)^p + \left(\frac{1-\lambda}{2}\right)^p \right)^{1/p} \tag{74}$$

Now consider a positive $2n \times 2n$ matrix M:

$$M = \begin{pmatrix} A & C \\ C^* & B \end{pmatrix} \tag{75}$$

The map $\Delta \otimes I$ acts on M via

$$(\Delta \otimes I)(M) = \begin{pmatrix} \lambda_+ A + \lambda_- B & \lambda C \\ \lambda C^* & \lambda_- A + \lambda_+ B \end{pmatrix} \tag{76}$$

Let $p \geq 2$, and let $q \leq 2$ be the index conjugate to p. Then as explained at the start of section 2, there is a positive $2n \times 2n$ matrix K satisfying $||K||_q = 1$ such that

$$||(\Delta \otimes I)(M)||_p = \mathrm{Tr}\Big(K(\Delta \otimes I)(M)\Big) \tag{77}$$

Following the methods used in (16), this leads to

$$\begin{aligned} \mathrm{Tr}\Big(K(\Delta \otimes I)(M)\Big) &\leq \left|\left| \begin{pmatrix} \lambda_+ ||A||_p + \lambda_- ||B||_p & \lambda ||C||_p \\ \lambda ||C||_p & \lambda_- ||A||_p + \lambda_+ ||B||_p \end{pmatrix} \right|\right|_p \\ &= ||\Delta(m)||_p \end{aligned} \tag{78}$$

where m is the 2×2 matrix

$$m = \begin{pmatrix} ||A||_p & ||C||_p \\ ||C||_p & ||B||_p \end{pmatrix} \tag{79}$$

By definition of the p-norm this implies

$$||(\Delta \otimes I)(M)||_p \leq \nu_p(\Delta) \Big(||A||_p + ||B||_p\Big) \tag{80}$$

Now let ρ be a $2n \times 2n$ density matrix,

$$\rho = \begin{pmatrix} \rho_{11} & \rho_{12} \\ \rho_{21} & \rho_{22} \end{pmatrix} \tag{81}$$

and consider the case where $M = (I \otimes \Phi)(\rho)$ and Φ is some other channel, so that $(\Delta \otimes I)(M) = (\Delta \otimes \Phi)(\rho)$. Then

$$A = \Phi(\rho_{11}), \quad B = \Phi(\rho_{22}) \tag{82}$$

and hence

$$||A||_p + ||B||_p \leq \nu_p(\Phi)\ \mathrm{Tr}(\rho_{11} + \rho_{22}) = \nu_p(\Phi) \tag{83}$$

Therefore (80) implies that

$$||(\Delta \otimes \Phi)(\rho)||_p \leq \nu_p(\Delta)\, \nu_p(\Phi) \tag{84}$$

Since (84) is valid for all ρ, we get

$$\nu_p(\Delta \otimes \Phi) \leq \nu_p(\Delta)\, \nu_p(\Phi) \tag{85}$$

and this establishes the result (72), since the inequality in the other direction follows by restricting to product states.

Acknowledgements This work was supported in part by National Science Foundation Grant DMS–0101205.

References

[1] G.G. Amosov, A.S. Holevo, and R.F. Werner, "On Some Additivity Problems in Quantum Information Theory", *Problems in Information Transmission*, **36**, 305 – 313 (2000).

[2] K. Ball, E. Carlen and E. Lieb, "Sharp uniform convexity and smoothness inequalities for trace norms", *Invent. math.* **115**, 463 – 482 (1994).

[3] C. H. Bennett and P.W. Shor, "Quantum Information Theory" *IEEE Trans. Info. Theory* **44**, 2724–2748 (1998).

[4] L. Gross, "Logarithmic Sobolev inequalities", *Am. Jour. Math.* **97**, 1061 – 1083 (1975).

[5] O. Hanner, "On the uniform convexity of L^p and l^p", *Ark. Math.* **3**, 239 – 244 (1958).

[6] C. King, "Maximization of capacity and l_p norms for some product channels", *Jour. Math. Phys.* **43**, no. 3, 1247 – 1260 (2002).

[7] C. King, "Additivity for unital qubit channels", *Jour. Math. Phys.* **43**, no. 10, 4641 – 4653 (2002).

[8] P. W. Shor, "Additivity of the classical capacity of entanglement-breaking quantum channels", *Jour. Math. Phys.* **43**, no. 9, 4334 – 4340 (2002).

[9] R. F. Werner and A. S. Holevo, "Counterexample to an additivity conjecture for output purity of quantum channels", *Jour. Math. Phys.* **43**, no. 9, 4353 – 4357 (2002).